حول التفسيرات المتباينة لنتائج الاختبارات

د. امينة محمد كاظم *

مقدمة

من أهم أهداف العلم ، الفهم والتنبؤ والتحكم . وفهم ظاهرة ما ، معناه اكتشاف العلاقة الوظيفية بينها وبين غيرها من الظواهر . أن معرفة شكل العلاقة وتقديرها تقديرا كميا دقيقا يستتبعه دقة في التنبؤ بحدوث الظاهرة ، ومن ثم التحكم الدقيق في أحداثها . وقد كان التقدم الهائل في العلوم الطبيعية يرجع بدرجة كبيرة الى أكتشاف الطرق والوسائل التي تساعد على فهم الظواهر الطبيعية فهما كميا بلغ من دقته أن أوصل الانسان الى أكتشافات كانت لعهد قريب محض خيال .

ولم يتوان العلماء والمفكرون في المجال السلوكي عن محاولة التوصل الى الطرق والوسائل التي تساعد على فهم وتفسير الظواهر السلوكية لعلهم في يوم يلحقون بمثل ما توصل اليه العلماء والمفكرون في مجال العلوم الطبيعية . ولم يمنعهم ذلك من التعاون مع غيرهم من العلماء والمفكرين في أغلب مجالات المعرفة للتوصل الى فهم أدق الظواهر السلوكية . . من هنا تعددت المحاولات والاتجاهات التي هي في الواقع اثراء للتفكير الإنساني الذي لم يقف مكتوف الايدي ، حتى يصل الى ما يريده ويبغيه ، وعندئذ تظهر في الافق علامات استفهام جديدة تجعله يلهث من جديد كي يتوصل للاجابة عليها .

أقترح Carver تصنيف أبعاد القياس السلوكي الى بعدين ، الأول وهو بعد القياس النفسي (السيكومتري) ويهتم بقياس الفروق بين الأفراد وهو الاهتمام التقليدي للاختبارات النفسية ، والآخر وهو بعد القياس التربوي (الايديومتري) ويهتم بقياس النمو أو الاكتساب عند الافراد وهو الاهتمام التقليدي للاختبارات التربوية . (١) وقد ابرز اتجاه جديد لا يهتم بذلك التصنيف ويهدف الى التوصل الى مفهوم جديد يحل مشكلة الموضوعية في تفسير نتائج الاختبارات ويقترب بذلك من المقاييس في العلوم الفيزيائية ، ويعتمد هذا الاتجاه الجديد على النماذج الاحتمالية (اي التي تقوم على نظرية الاحتمالات) ومن أمثلتها نموذج راش .

---

*استاذ علم النفس المساعد بجامعتي عين شمس والكويت

وتهدف هذه الورقة الى محاولة التوفيق بين هذه الاتجاهات الثلاث المعاصرة في تفسير نتائج الاختبارات . ويعتبر هذا الاتجاه الأخير هو أحدثها جميعاً لذا سيحظى اكثر من غيره من الاتجاهـات بشيء من التفصيل .

بعد القياس النفسي (البعد السيكومتري)
Psychomrtric Dimension

أشار Cronbach (١٩٧١) الى ان حركة القياس في علـم النفس قد بدأت مع تأكيدات داروين على الفروق بين الافراد ، وأن كل الجهود النظرية وراء الدرجات الاختبارية قد بذلت بهذا المفهوم . وعلى هذا فقد كان الهدف التقليدي للمقاييس النفسية هو قياس الفروق الفردية ، سواء في القدرات العقلية او في باقي سمات الشخصية . وقد اطلق على هذه المقاييس اسم المقاييس جماعية المرجع (Norm - Referenced Measurements) حيث لا يكون لدرجاتهـا الخـام معنى ، ما لم ترد أو تقارن بدرجة متوسطة ، وذلك اما بالانحراف عن المتوسط او بواسطة المئينيات المحسوبة لدرجات مجموعة اختبارية . وقد اطلق Carver على هذا البعد ،بعد القياس النفسي (٢)

وقد أعلن Popham أن هناك كثيرا من الأسباب التي تجعـل تلك الاختبـارات التقليدية (جماعية ـ المرجع) غير مفيدة عمليا في مجال التقويم التربوي ، ثم أشار الى المقـاييس (محكية ـ المرجـع)(Critrerion Referenced Measurements) كأتجاه جديد أكثـر ملاءمـة لمتطلبـات التقويم التربوي .(٣)

نقد المقاييس النفسية :

سنتناول في نقدنا للمقاييس على بعد القياس النفسي ناحيتين :

- مناقشة بعض المفاهيم الاساسية للمقاييس التقليدية (جماعية ـ المرجع) .

- مناقشةمدى اغفال المقاييس النفسية التقليدية لبعض أهداف القياس الهامة .

ويمـكن تلخيص الاعتراضـات الـرئيسية على المقـاييس التقليدية للقـدرات ومصطلحاتها ، نسبة ذكاء ، عمر عقلي . . . الخ فيا يلي :

١ ـ الاعتاد على معيار جماعة معينة في القياس .

٢ ـ عدم وجود مقياس قدرة خاص (مثل مقياس الوزن أو الطول) ، يمكن تطبيقه على المدى الواسع من القدرة .

٣ ـ مشكلات الثبات التي تواجه غالبا عند عمل التدريج المعياري للافراد الذين يتميزون بمستوى متطرف من القدرة . (٤)

وسنقف عند هذه الاعتراضات واحدا بعد الآخر .

النقطة الأولى :

أ ـ مفهوم العمر العقلي :

أن مفهوم العمر العقلي ليس مفهوما مقنعا (أو كافيا) ، وذلك لاعتماده على المقارنة بجماعة ـ مرجعية . أنه لمايثير العجب حقا لو كان التعبير عن طول الطفل أو وزنه ، بمصطلح الطول المتوسط أو الوزن المتوسط لأقرانه من الأطفال عند سن معينة ، او لو كان التعبير عن درجة الحرارة في أحد الأيام من شهر معين بمتوسط درجات الحرارة في ذلك الشهر . إن التعبير عن مستوى القدرة يكون أكثر معنى اذا عبر عنه بمقياس ذي وحدات ، مثل القدم أو البوصة أو الرطل أو الكيلو أو درجة الحرارة .

ب ـ أدوات القياس :

لا يصح أن تعتمد المقاييس الموضوعية على أداة القياس المعينة التي استخدمت بافتراض دقة هذه الأدوات . أن المستقيم الذي طوله ٥ سم هو دائما ٥ سم بصرف النظر عن المسطرة التي استخدمت في القياس ، ومهما كان الحد الذي ينتهي عنده تدريجها ، ومن أية نقطة من تدريجات المسطرة بدأ القياس . كما لا يعنينا ما هي تلك المستقيات الأخرى التي استخدمت تلك المسطرة في قياسها ، وهذا ما يعني بأداة القياس الموضوعية التي لا تتأثر بمعيار الجماعة . Norm - free) objective measurement) .

جـ ـ وحدة القياس :

نلاحظ أن هدف القياس في علوم الفيزياء هو ايجاد مقاييس نسبة (Ratio Scales) مثل قياس الطول بالأمتار ، أو مقاييس الوحدات المتساوية (ingennal cales) مثل قياس درجات الحرارة بالدرجات المئوية أو الفهرنهيتية ، حيث يكون للتدريج وحدة القياس نفسها ، سواء كنا نقيس درجة حرارة سائل ما أو درجة أنصهار مادة ما ، فان التغير بمقدار درجة مئوية واحدة له المعنى نفسه في الحالتين ، ولكن ينبغي أن ندرك أن حجم وحدة القياس ، وكذلك نقطة الصفر هو أمر اصطلاحي . ولهذا فمن المسموح به القيام بأي تحويل خطي لايجاد تدريج آخر ذي وحدات متساوية وله الخواص نفسها ، مثل تحويل مقياس الحرارة الفهرنهيتي الى المئوي أو العكس ، وذلك اذا كان مدى التدريج غير ملائم من وجه النظر التطبيقية .

قد تكون مقاييس النسبة المطلقة للظواهر السلوكية الواسعة المدى ، بما في ذلك القدرات ، هدفا غير واقعي ، وذلك لاستحالة تفسير معنى المستوى الصفرى (لقدرة ما مثلا) . وعلى كل حال فان ما نتوق اليه فعلا هو ايجاد مقاييس موضوعية تستخدم فيها وحدة قياس واحدة عند كل مستوى من مستويات الظاهرة السلوكية .

ومع ذلك فأن المفاهيم المنتسبة لمعايير الجماعة مثل نسبة الذكاء الانحرافية تكون مفيدة اذا كان المطلوب هو مجرد مقارنة أداء فرد على اختبار ما مع أداء أقرانه .

النقطة الثانية :

أن نقطة الاعتراض الثانية هي نتيجة مباشرة للأولى ، وقد غطى معظمها أثناء تناول النقطة الأولى ، حيث يتضح أيضا أن المقاييس التقليدية المعتمدة على معايير الجماعة لا تتيح المقارنة المباشرة المتطلبة في دراسات النمو المستعرضة أو الطولية . فمن غير المستطاع قياس النمو النموذجي (typical) لتحصيل ما أو قدرة ما عبر العمر ما دامت قياسات كل فرد يعبر عنها بواسطة المستوى المتوسط لعينة عمره .

النقطة الثالثة :

تختص هذه النقطة ببعض المشاكل الناتجة عن المستويات المتطرفة من الظاهرة السلوكية ، فاذا فرضنا أن طفلا مفرط الذكاء في السادسة من العمر ، استطاع أن يجيب عن جميع بنود اختبار مقنن على مجموعة من الأقران من العمر نفسه ، عندئذ ينبغي الاعتماد على اختبار آخر يكون أكثر مناسبة له ، ومن ثم يكون مقننا على مجموعة من الأطفال أكبر عمرا ، ولنفرض أنهم في سن التاسعة مثلا . وعندئذ فأن أداء هذا الطفل على هذا الاختبار الأخير يقارن بمعايير عمر التسع سنوات وقد يحدد له عمر عقلي على هذا الاساس . ولكن هل من الممكن عندئذ أن يكون التعبير بمفهوم نسبة الذكاء له معنى ؟ لكي يكون للتعبير معنى ، ينبغي أن يقارن أداء هذا الطفل على هذا الاختبار الأخير مع أداء غيره من اطفال سن السادسة على هذا الاختبار نفسه . ولكن واقع الأمر أنه ليست هناك أية معلومات عن الأداء المتوقع لهؤلاء الأقران من سن السادسة على هذا الاختبار الخاص بسن التاسعة .

أما من حيث أغفال المقاييس النفسية التقليدية لبعض أهداف القياس الهامة ، فلنا أن نتساءل :

هل من الصحيح أنه لا يعنينا في مجال القياس النفسي سوى مقارنة أداء الأفراد بضعهم ببعض وأنه لا يعنينا مقارنة أداء الفرد بمستوى أداء معين ، أو بمستوى أداء الفرد نفسه في مواقف سيكلوجية مختلفة ، أو بعد فترة زمنية معينة ؟

قد يكون من المفيد أن نورد بعض الأمثلة من بعض المجالات في علم النفس :

مقارنة أداء الفرد بمحك أو مستوى معين من الأداء :

عندما يتقدم عدد من الطلبة للدراسة بأحدى كليات الجامعة وليس لدى الجامعة أي محك للاختبار سوى معيار الجماعة فتضطر عندئذ الى قبول أحسن مجموعة من هؤلاء المتقدمين تبعا لمعاييرهم الخاصة على أداة القياس ، حتى تصل الى عدد الطلبة المناسب لإمكانياتها المتاحة . فماذا تكون النتيجة ؟

لو فرضنا صدق أداة القياس في قياس المقومات اللازمة للنجاح في هذه الكلية ، أليس من المحتمل ألا يتوافر القدر اللازم من هذه المقومات حتى عند هؤلاء الطلبة ذوي المستويات المرتفعة بالنسبة لباقي أقرانهم المتقدمين لهذه الكلية ؟ أن معايير الجماعة على صفه ما ليست ثابتة بل تتغير مستوياتها من مجتمع لآخر . هنا تبدو أهمية تحديد المستوى أو المحك المناسب الذي ينبغي ان يصل اليه الطالب كي يمكنه متابعة الدراسة بنجاح في هذه الكلية . وتبدو أيضا أهمية تحديد المحك أو المستوى المعين من الأداء عندما يتطلب الأمر اختيار الأفراد الصالحين لعمل معين يتطلب مستوى معين من المهارة (مثل أداء الطيارين) ، أو مستوى معين من صفة أو صفات سلوكية معينة (مثل متطلبات النجاح في مهنة ما) .

هنا تبرز مشكلة تحديد المحك المناسب لأهداف القياس ، ومن الذي يقوم بتحديده ؟ وعندما نكون بصدد تحديد المحك في اختبارات الاتجاهات والقيم هل نستطيع أغفال المستوى الحضاري والثقافي السائد في المجتمع وهو ما يختلف من مجتمع لآخر ؟

ستتعرض لنقطة تحديد المحك عند مناقشة المقاييس محكية ـ المرجع .

مقارنة أداء الأفراد بأدائهم السابق :

ويبدو هذا في مجال دراسة تغير الاتجاهات والقيم وأحيانا في مجال قياس نمو القدرات حيث يقاس التغير في اتجاهات أو قيم ، أو أداء أو سلوك الأفراد بعد فترة زمنية معينة أو بعد برنامج معين لتعديل الاتجاهات أو القيم أو السلوك أو في مواقف سلوكية مختلفة .

ترى هل تحقق المقاييس النفسية بوضعها الراهن وفلسفتها الراهنة هذا الغرض بمعنى هل من الممكن أن تقيس مدى النمو أو الاكتساب أو التغير في صفه سلوكية ما ؟ ينبغي من وجهة نظر القياس النفسي أن تكون مفردات الاختبارات والمقاييس حساسة للفروق الفردية لذا تعتبر أحسن البنود هي تلك التي يصل مقدار تباين درجاتها الى نهايته العظمى (**٢٥ و**) ، وذلك عندما يتساوى كل من معامل السهولة ومعامل الصعوبة لهذا البند ويصبح كل منهما (**٥ و**) وعلى هذا تستبعد تلك المفردات التي تساوي معاملات سهولتها الصفر أو الواحد الصحيح ، وتستبقي تلك التي تقترب معاملات سهولتها من **٥ و** . على قدر الامكان . ولكن عندما تكون بصدد قياس اتجاهات مجموعة من الأفراد ، فلنا أن نتساءل ، أليس اجماع مجموعة من الأفراد على وجهة نظر معينة ازاء قضية معينة له معنى يدل عليه ؟ فلو أجابت عينة من الأفراد مثلا بالنفي على سؤال عن رأيها في تأييده فكرة تنظيم الأسرة ثم بعد برنامج تثقيفي معين كانت أجابات العينة كلها على السؤال نفسه بالايجاب والتأييد ، أليس في هذا الدليل على فاعلية ذلك البرنامج في تغيير اتجاه العينة من المعارضة التامة لفكرة تنظيم الاسرة الى التعضيد التام . ان قياس هذا التغيير في الاتجاه لا يقل أهمية عن قياس الفروق في الاتجاهات بين هؤلاء الأفراد أن لم يكن أهم في أحوال كثيرة .

وهناك أمثلة أخرى تبدو فيها أهمية مقارنة أداء الأفراد بأدائهم السابق مثل التغيير في حالة

القلق في مواقف مختلفة الشدة والتعصيب وعلى هذا قد يكون من المنطقي الابقاء على تلك المفردات ذات معاملات السهولة المتطرفة (صفر أو ١) ، وهذا ما أغفلته المقاييس النفسية التقليدية بل واعتبرته من العيوب التي ينبغي تلافيها للوصول الى أداة القياس السليمة .

مقارنة أداء الفرد بامكانيات ذاته :

عندما لا يكون الهدف هو الإنتقاء ينبغي أن نتحرر من سيطرة فكرة مقارنة الفرد بغيره من الأفراد بحيث لا تكلف نفس الا وسعها . عندئذ نستطيع توجيه الأفراد كل حسب أمكانياته العقلية والشخصية بحيث يعمل على تنمية اكتسابه لأساليب السلوك بالقدر الذي تؤهله له امكانياته الحقيقية بصرف النظر عما يكون عليه وضعه بالنسبة للأقران . هنا تبدو أهمية ابتكار الطرق والوسائل التي تمكننا من عمل هذه المقارنات بين السلوك الواقعي للفرد وبين أفضل ما يمكن أن يعطيه . ويتبع ذلك وضع الأساليب اللازمة لمعالجة الفجوة ـ ان وجدت ـ بين سلوك الفرد الواقعي وبين امكانيات ذاته حتى تصل بالفرد الى استثمار جميع امكانياته فيفيد نفسه ومجتمعه ، وهذا من أهداف التوافق الهامة .

بعد القياس التربوي (الايديومتري)
Edumetric Dimension

كان البعد الآخر من أبعاد القياس السلوكي هو ما أطلق عليه(Carver) بعد القياس التربوي (الايديومتري) وفيه يجسد ما أغفلته النظرة السيكومتريه التي أشرنا اليها آنفا ، وحيث يكون تقويم الاختبار بمصطلحات القياس التربوي . أي الى أي مدى تعكس النمو الفردي (Within indvidual growth) الذي كان الاهتمام التقليدي الأول للاختبارات التربوية . خاصة بعد التغير الذي حدث في مفهوم العملية التعليمية . فبعد أن كان اهتمام المدارس منصباً على الانتقاء والتنبؤ بالنجاح بمعنى تقديم الخدمة لأعداد قليلة ممن هم أفضل ، أصبح التعليم حقًا للجميع ولم يعد السؤال عن الاختيار والتنبؤ بالنجاح بقدر ما أصبح كيف نعلم الجميع بحيث يحصل كل فرد على أفضل ما يمكن ان يعطيه لنفسه ومجتمعه بصرف النظر عن وضعه بين الاقران .

وظهر مفهوم التعليم من أجل الاتقان (٥) .وظهرت المناداة بالابتعاد عن شكل الناقوس الذي يميز التوزيع الاعتدالي أو ما يسميه البعض المنحنى الطبيعي . وفي هذا يقول بلومBloom «ليس هناك شيء مقدس بالنسة للمنحنى الاعتدالي ، فهو التوزيع الاكثر مناسبة للنشاط القائم على الصدفة أو العشوائية . ولكن التربية نشاط مقصود ونحن نسعى لكي نجعل الطلبة يتعلمون ما ندرسه . فاذا كنا فعّالين في تعليمنا فان توزيع درجات التحصيل ينبغي ان يكون مختلفا جدا عن

المنحنى الاعتدالي . وفي الحقيقة قد نصل الى الحد الذي نقرر فيه ان جهدنا التربوي يكون فاشلا بقدر اقتراب توزيع درجات التحصيل من التوزيع الاعتدالي» . وعلى هذا اصبح الغرض الاول لاختبارات القياس التربوي ليس التركيز أساسا على الفروق بين الافراد بل قياس مدى الاكتساب او النمو للفرد في نواح مختلفة مثل المعلومات او المهارات او التحصيل . ومن ثم فهي لا تحدد مستوى الفرد بالنسبة لمعايير أقرانه ، بل تهتم بمقارنة الفرد بالنسبة لميزان أو محك للاداء يحدد حسب الاهداف الموضوعية للاختبار وبصرف النظر عن مستوى الجماعة . لذا فان تلك الاختبارات التي سميت بالاختبارات محكية المرجع Criterion refeernced tests والتي جعلها Carver مرادفة للمقاييس التربوية ، تختلف عن تلك التي ترد الى معيار الجماعة ، ليس فقط من حيث الهدف بل ايضا من حيث البناء ، فهي تختلف من حيث اختيار البنود ومستويات صعوباتها ، ومن حيث شروط صدقها وثباتها وتقنينها . فعلى سبيل المثال ، فان احسن البنود من وجهة النظر السيكومترية هي تلك التي يكون معامل سهولتها ٥ ر وأسوأها هي تلك التي يكون معامل سهولتها صفرا أو واحدا صحيحا . والعكس صحيح بالنسبة لوجهة النظر الأيديومترية ، حيث تكون أحسن البنود هي تلك التي يكون معامل سهولتها صفرا قبل تنفيذ البرنامج التربوي ثم يصبح واحدا صحيحا بعد تنفيذه . كذلك فان مفهوم الثبات من وجهة النظر السيكومترية تتعلق بثبات كفاءة الاختبار على التمييز بين الافراد ، في حين ان مفهوم الثبات من وجهة النظر الأيديومترية تتعلق بثبات كفاءة الاختبار في قياس الاكتساب او النمو لدى الافراد .

نقد المقاييس التربوية (الأيديومترية) :

سنتناول مناقشتنا للمقاييس التربوية ناحيتين :

الاولى : مناقشة بعض المشكلات الأساسية للمقاييس التربوية (محكية ـ المرجع)

الثانية : مناقشة مدى أغفال تلك المقاييس لبعض الاهداف الهامة في القياس .

وقد يكون من أهم المشكلات الخاصة بالمقاييس محكية ـ المرجع ما يأتي : -

١ ـ مشكلة اختيار المحك المناسب للأداء :

أوضح (Ebel , R . L) في مناقشته لحدود المقاييس محكية ـ المرجع «أنه من الصعب التوصل الى تلك المقاييس على أساس متين ، فمن الصعب التوصل الى المحك الصادق الذي تبنى عليه هذه المقاييس ، وهذا يرجع الى عدم الاتفاق على المفاهيم ، القيم ، ومستويات المدرسين . ومحاولة توضيح الأهداف انما هو عملية استهلاك وقت» . (٧)

ونحن نتساءل بدورنا اذا كان المحك هو مستوى أداء معين فكيف يحدد هذا المستوى ؟ ومن الذي يحدده ؟ هل يحدده العلماء المتخصصون في المجال ؟ أم يحدده التربويون ؟ أم علماء النفس ؟ أم المخططون لسياسة التعليم والمجتمع ؟ أم يحدده علماء الاحصاء ؟ أو المدرسون ؟ أم نعود الى معيار الجماعة كمحك للأداء ؟ وما هو الحال بالنسبة للاختبارات التحصيلية حيث تختلف طرق التدريس ومستويات المدرسين ؟ مما يجعل تحديد محك واضح للنجاح على الاختبار أمرا صعبا للغاية ، وظهرت طرق وابتكارات جديدة لعلها توصل الى المعيار المناسب لتحديد مستوى الأداء المطلوب ومن بينها تلك المعايير العملية التي من أمثلتها المعيار الذي وضعه(William Gray) لمحو أمية الأفراد وهو ، عندما يتقن الفرد القراءة والكتابة والحساب . . . . الى مستوى لا يجعله يرتد مرة أخرى وجعل هذا المعيار هو مستوى انتهاء الفرد من المرحلة الابتدائية .

٢ ـ مشكلة الصدق :

يتصل الصدق الاجرائي للاختبار الأيديومتري بمدى حساسيته للاكتساب أو النمو فاذا أظهر الاختبار أكتسابا أو نموا قليلا حيث يتوقع اكتساب أو نمو كبير فهذا دليل على ضعف صدق الاختبار . ولا يمكن عادة تحديد صدق الاختبار الأيديومتري باجرائه على مجموعة من الأفراد في وقت واحد كما هو الحال في الاختبارات السيكومترية ، انما يكون عادة اجراء الاختيار في موقفين حيث يتوقع بينهما نمو أو اكتساب ، ثم يقارن الاكتساب الواقعي الحادث مع الاكتساب المتوقع لمعرفة صدق الاختبار .(٨)

هنا تبرز مشكلة الاكتساب المتوقع وكيف يحدد هذا التوقع ؟ وهل نعود الى معيار الجماعة لتحديد مستوى هذا الاكتساب المتوقع ؟

٣ ـ مشكلة الثبات :

يتعلق الثبات الايدومتري بثبات الاكتساب أو النمو عند الافراد كما ينعكس من الاختبار . وهناك طريقة جيده لقياس ثبات الاختبار الايدومتري وهو اجراء صور متبادلة من الاختبار في الاحوال قبل وبعد تعرض الافراد لبرنامج تربوي أو تثقيفي معين ، هنا يعطينا ثبات الاكتساب أو التغيير بين هذه الصور مقياساً جيدا لثبات الاختبار .

أن تلك المشاكل المتعلقة بقياس التغير والاكتساب تتضمن دائما مشاكل أرتباطية سيكومترية . يعرف مثلا كرنباخ وفرباي(Cronbach and Furby) ثبات الاكتساب أو درجات الفرق بأنها ارتباط الدرجة بفرق ملاحظ مستقل . يكون الاختبار الايدومتري ثابتا تمام الثبات عندما يفشل كل فرد قبل البرنامج ثم ينجح كل واحد منهم بعد البرنامج وعندئذ يكون ارتباط درجات الاكتساب بأي متغير آخر مساويا للصفر . وهنا يعتبر الاكتساب غير ثابت أطلاقا حسب نموذج كرنباخ وفرباي السيكومتري . يقول بررتر Brerter (١٩٦٣)«أنه لم يكن ليسمع أبدا من زملائه من يقبل أن يهمل أبحاثا ذات أهداف رئيسية لمجرد صعوبة المشاكل الاحصائية ، الا فيما

يتعلق بمشاكل قياس التغير .»(٩)

أما من حيث أغفال تلك المقاييس الايدومترية لبعض أهداف القياس الهامة فأهمها :

اغفال مقارنة أداء الفرد بأداء أقرانه :

من بين الاعتراضات التي وجهها(Ebel . R . L) في مناقشته لحدود المقاييس محكية ـ المرجع «أنها تنبؤ نا عن كل أو حتى عن معظم النواحي التي نحتاج لمعرفتها عن التحصيل التربوي ، أي عن التفوق أو القصور النسبي» وكما ذكر سابقا فإن هذه الاختبارات الايديومترية لا تعتمد في بنائها على تلك البنود التي تميز بين الافراد في الاداء،ولا تعتمد في تفسير درجة الفرد على الاختبار . على أدائه النسبي مقابل أداء المجموعة الاختبارية،ومن ثم لا نستطيع تحديد قصوره أو تفوقه النسبي والذي يعتبر من أهداف التقويم التربوي الهامة .

ويلخص(Ebel ,R , L) أعتراضاته في النهاية بقوله «أن المقاييس التي ترد الى محك ليست بالجديدة ، ولا نستطيع أن نتوقع أن تحسن بصورة فعالة تقويمنا التربوي . (١٠)

وقفة لامعان النظر :

قد يكون من المفيد أمعان النظر في ذلك التصنيف الثنائي الذي أقترحه(Carver) لأنواع المقاييس السلوكية ونقف قليلا لنتساءل هل تقتصر المقاييس السلوكية على هذين النوعين فقط ، جماعية ـ المرجع وهي ما جعلها(Carver) مرادفة للبعد السيكومتري ، محكية ـ المرجع وهي ما جعلها(Carver) أحد مرادفات البعد الايديومتري .

أن مقارنة أداء الفرد بغيره من الافراد أو مقارنته بمحك أو مستوى معين من الاداء ، أو قياس التغير الحادث في سلوك الافراد ، أو مقارنة أداء الفرد الواقعي بما يمكن أن تؤ هله له أمكانياته الذاتية ، كلها أهداف هامة سواء في المجال النفسي أو المجال التربوي .

لذلك قد يمكننا أن نعيد التصنيف بطريقة اخرى تبعا لأهداف كلا البعدين السيكومتري والايديومتري على النحو الاتي :

١) مقارنة أداء الفرد بمعايير الجماعة .

٢) مقارنة أداء الفرد بمحك أو مستوى معين من الاداء .

٣) مقارنة أداء الفرد بأدائه السابق .

٤) مقارنة أداء الفرد الواقعي بامكانيات ذاته .

٥) غير ذلك .

أن اختلاف المسلمات الواضح يجعل التوفيق عسيرا بين هذه الاهداف . وقد كان التركيز الاكبر قد جرى بصورة تقليدية على بعد القياس النفسي حتى أن كثيرا من الاختبارات التي قننت بهذا المفهوم قد أستخدمت لقياس الاكتساب أو النمو دون أن تكون قد طورت أو قومت من وجهة نظر القياس التربوي ، لذا برز الخطر حيث قد تكون هذه الاختبارات غير حساسة لقياس الاكتساب حين يكون هناك في الواقع نمو أو اكتساب .

ترى هل لم يعد أمامنا سوى أن نطبق المقاييس جماعية المرجع اذا كان الهدف هو مقارنة أداء الافراد بعضهم ببعض ؟ وأن تطبق المقاييس محكية المرجع عندما يكون الهدف هو مقارنة أداء الافراد بمستوى معين من الاداء ؟ وهلم جرا .

قد يبدو هذا معقولا خاصة أن فتح النوافذ لجميع الاتجاهات والافكار تجعل الرؤ ية متعددة الجوانب والزوايا . ولكن حدثت المحاولات الجديدة لابتكار طرق القياس التي قد تحل مشكلة تعدد الاهداف وأختلاف المسلمات وما يستتبع ذلك من تعدد الطرق والوسائل للوصول الى هذه الاهداف . . . وقد كان الامل الذي راود الكثيرين هو الوصول الى مقاييس تشبه تلك المقاييس الخاصة بالمجالات الفزيائية وعندها قد نصل الى الموضوعية التي قد تحقق جميع أهداف القياس ولنضرب لذلك مثلا :

اذا كان طول فتى في سن الثامنة عشر هو ١٨٥ سم فهو :

- أطول من متوسط طول أقرانه (جماعية - المرجع) .
- وصل وتعدى مستوى الطول المطلوب في الكليات العسكرية (محكية - المرجع) .
- ازداد طوله عما كان عليه في سن الخامسة عشر (مقارنة طوله بطوله السابق) .
- اقصر من امكانيات الطول التي تحددها عوامله الوراثية (مقارنته بامكانيات ذاته) .
- أقصر مما يود أن يكون (مقارنته بمستوى طموحه) .

وعلى هذا فان تقدير طول الفتى بوحدة قياس مطلقة أمكننا من أن نحقق الاهداف المختلفة للقياس وعلى هذا أبرزت الحاجة لمقياس موضوعي لقياس السلوك ، يشبه تلك المقاييس المستخدمة في العلوم الطبيعية ، سواء كان ذلك على بعد القياس النفسي أو بعد القياس التربوي . وهذا يتطلب أولا أن تكون مواصفات بنود الاختبار بحيث لا تعتمد على توزيع أداء المجموعة التي أجرى عليها الاختبار ، وثانيا أن يعطى الاختبار تقديرا للأداء لا يعتمد على مجموعة البنود المعينة التي تكون الاختبار . وقد حدد(Murray , D ..) مطالب الاستقلالية والموضوعية في التدريج لكل من صعوبة البند وقدرة الفرد ، ويمكن تلخيصها تحت أربع نقاط :

١) لا تعتمد صعوبة البند على قدرة الافراد الذين يجيبون عنه .

٢) لا تعتمد صعوبة البند على باقي بنود الاختبار .

٣) لا يعتمد تقدير الفرد على البنود المعينة التي يجيب عنها .

٤) لا يعتمد تقدير قدرة الفرد على قدرة باقي الافراد الذين يجيبون على الاختبار .

تحت هذه الشروط فقط يرى (Murray) أنه من الممكن أن نحصل على نوع من المقاييس الموضوعية التي نصبو اليها . (١١)

وقد كانت النماذج الرياضية القائمة على نظرية الاحتمالات من أهم الوسائل التي ساعدت على السير نحو تحقيق حلم الموضوعية في القياس السلوكي وان كان الطريق لا زال طويلا ولا زالت تكتنفه الصعوبات .

النماذج الرياضية :

اتسع مجال استخدام الرياضيات في مختلف ميادين الحياة ، وأمتد حتى شمل أغلب ظواهرها سواء الاجتماعية ، أو الاقتصادية أو الفيزيائية* . وأصبح الاتجاه الجديد هو ترجمة ظواهر الحياة المختلفة الى صيغ رياضية مناسبة وهو ما نسميه بالنمذجة ، حيث يبدأ بالمشكلة الحقيقية من الواقع أي من الظواهر الواقعية من مختلف ميادين الحياة ، ثم ترجمة هذه الظواهر الى نماذج واسطة لتوضيح المتغيرات المؤثرة في الظاهرة ثم تحويلها الى نماذج رياضية بحتة حيث يمكن دراستها وحلها بغض النظر عن معناها الاصلي - ثم أرجاعها وأستخدامها وتطبيقها على الظاهرة الاصلية .(١٢)

ومن بين النماذج الرياضية التي أستخدمت في المجال السلوكي تلك النماذج الاحتمالية والتي كان من أبرزها نموذج راش والذي سوف نتعرض له بشيء من التفصيل نظرا لجدة هذا الاتجاه ، حيث نبين أهدافه ومسلماته ومبادئ القياس التي يرتكن اليها وكيفية تقدير أبعاده ثم مناقشته من حيث تحقيقه للموضوعية مع توضيح وتطبيق هذا النموذج على أحد الاختبارات ، وكيفية الاستفادة منه في البيئة المصرية وفي جامعة الكويت ثم محاولة مناقشة هذا النموذج مناقشة نقدية من حيث هل تحقق الهدف الذي وضع لأجله فعلا أم لا زال الشوط طويلا .

نموذج راش
The Rasch Model

أفترضه عالم الرياضيات الدانمركي «جورج راش» حتى يمكن التوصل الى الموضوعية وذلك خلال الستينات ، ومع ذلك فان تطبيقاته العملية في مجال القياس النفسي ظلت قليلة حتى سنوات قريبة . (١٣)

وفي السنوات الاخيرة . طور «بن رايت» ، الاستاذ بجامعة شيكاغو ، هذا النموذج من

* تستخدم هذه النماذج أيضا في حل مشاكل المرور وأنشاء المدن وشبكات المواصلات والتليفونات . . . . الخ .

الوجهة العملية وفي عام ١٩٧٣ طبق نموذج «راش» في انجلترا على نتائج التحصيل المدرسي على يد (Alan Willmot and Diana Fowles) وذلك في أحد مشروعات المركز القومي للبحوث التربوية N . F . E . R كما أستخدم أيضا هذا النموذج على نطاق واسع في عمل مقياس قومي للذكاء في بريطانيا B . I . S وذلك بداية من عام ١٩٦٥ ، وحتى عام ١٩٧٧ لم يكن قد فرغ منه بعد . ولم يكن في بناء وتقنين أختبارات هذا المقياس ، على أساس نموذج راش ، ولا في تفسير درجات الافراد باصطلاحات هذا النموذج ما يمنع من التحليل وما يتبعه من التقنين بمصطلحات المقاييس التقليدية جماعية ـ المرجع(Murray . D ., 1976 . P . 41) . (١٤) .

وتهدف هذه النظرية الجديدة الى الوصول الى مقاييس ، لا تعتمد على أداة قياس معينة ولا على توزيع أداء المجموع الاختبارية ، الى الوصول الى الموضوعية في القياس السلوكي . وسنبدأ هنا بأفتراضين عرضهما (Murray) وهما يتناولان بنود الاختبار والفرد الذي يجيب عن هذه البنود . وعلى الرغم مما يبدو من بساطة هذين الفرضين فلا يستتبع ذلك أن يكون أستيفاؤهما سهلا . (١٥)

الفرض الاول :

أن بنود الاختبار ذات صعوبة أحادية البعد ، ولا تعتمد هذه الصعوبة على صعوبات البنود الاخرى بالمجموعة أو على قدرة الافراد الذين يجيبون عنها وهذا يعني أنه الى الحد الذي تميز فيه البنود بين الافراد ذوي المستويات المختلفة من قدرة ما ، فأن هذه البنود جميعها تميز بين الافراد على أمتداد البعد لهذه القدرة دون غيره . أما استقلال مستوى صعوبة البنود عن بعضها وعن الافراد الذين يجرونها ، فهذا يعني أنه اذا قدر مستوى صعوبة بند ما بمقدار ضعف صعوبة بالنسبة لبند آخر ، فأن صعوبته تكون دائما الضعف بالنسبة لأي فرد عند أي مستوى من القدرة يؤدي هذين البندين ، دون النظر بأي من البنود قدرت قدرة هذا الفرد . وهذا يعني أن جميع هذه البنود ينبغي ان يكون لها قدرة تمييز متساوية ، لذا فهي تحدد بواسطة بعد واحد(Single Parameter) هو صعوبتها .

الفرض الثاني :

أن الافراد ذوي قدرة أحادية البعد ، تحدد وحدها مستوى أدائهم على اختبار ما ، ولا تعتمد هذه القدرة على قدرة أية مجموعة أخرى من الافراد ممن يؤدون الاختبار نفسه ، أو على صعوبات البنود التي يؤدونها .

بعبارة أخرى ، طالما أختلف الناس في أدائهم على بنود أختبار ما (على الاقل على مستوى الاحتمال) فأن أداءهم على كل بند يحدد ببعد واحد للقدرة(Single ability Parameter) واذا فرضنا أن صعوبات كل البنود التي اداها الفرد معروفة ، فان بعد القدرة هذا يمكن حسابه ويكون ثابتا للفرد نفسه بصرف النظر عن أي اختبار فرعي ـ ذي البنود المتدرجة ـ قد أستخدم .

ولهذين الفرضين معا معنى ، هو أن كلا من بنود الاختبار والافراد الذين يؤ دون الاختبار قد وصف بمقياس واحد ، وأن كل صعوبات البند تقيس نفس الشيء مثل بعضها البعض ومثل ما تقيسه قدرات الفرد . وهذا يعني أن مقاييس البند والفرد تكون على نفس التدريج .

أن أساس النموذج هو ، أنه من الممكن التحدث عن مجابهة(encounter) الفرد للبند ، وهو في هذه المجابهة أما أن يكون ناجحا أو فاشلا ، أي يكون على صواب أو خطأ ، ويعبر عن ذلك بأصطلاحات عاملين فقط أحدهما خاص بالفرد والاخر خاص بالبند . ويشكل هذان العاملين بعديParameters النموذج ، بحيث يكون كل منهما ثابتا لفرد معين أو لبند معين . فيكون البعد الخاص بالفرد مقدارا ثابتا لفرد معين بصرف النظر عن أي البنود أداها الفرد ، ويكون البعد الخاص بالبند مقدارا ثابتا لبند معين بصرف النظر عن أي من الافراد أدى هذا البند .(١٦)

فاذا أمكن عمليا توفير هذين الشرطين من الممكن التوصل الى الموضوعية . وعندئذ لا يهم أي مجموعة من البنود الاختبارية يؤ ديها الفرد في فترة ما ، كما لا يهم من من الافراد يؤ دي على مجموعة معينة من البنود الاختبارية .

مبادىء القياس :

قبل أن نناقش كيف يعمل النموذج ، نقدم مبدأين من مبادئ القياس قام. Willomt , A and Fowles D ., بأقتباس تصويرهما من العلوم الطبيعية . اذا كنا نريد قياس طول شخص معين ، وليس لدينا أية أداة للقياس سوى حزمه من العصي كل واحدة محدد عليها الطول الخاص بها . فاذا قورنت كل عصا بطول هذا الشخص فأن نجاحه في تخطي طول العصا أو فشله في تخطيها يمكن تحديده بالعبارة نجح/ فشل تقال عند كل مقارنة بناء على كون الشخص أطول من العصاأو أن العصا هي الاطول . وبالطبع هناك بعض العصى التي تحير فعلا هل تخطاها الفرد أم لا . هذه العصى الاخيرة المحيرة التي تعطينا عددا متساويا من الاحكام نجح/ فشل (أي ٥٠٪ لكل حكم) هي التي يكون طولها هو أحسن تقدير لطول هذا الفرد .

أن أول مبدأ للقياس هنا هو أن تكون أداة القياس المستعملة مناسبة للصفة المقاسة ، فالساعة مثلا غير صالحة لقياس الاطوال . فلا يعنينا شكل الادارة ما دامت مناسبة لقياس الصفة (مسطرة شريط مدرج . .) . كما لا يعنينا أيضا الاختلاف في نوع التدريج ما دام مناسبا لتقدير الصفة . ففي مثالنا السابق قد يكون طول العصا المناسبة (٥٠٪ نجح/ فشل) هو قدم بوصة ٥ ، فأذا أعيدت التجربة بواسطة مجموعة أخرى من العصي فأن طول العصا المناسبة للفرد نفسه من الممكن أن يكون ١٦٥ سم .

أما المبدأ الثاني : فهو أن قياس صفة متصلة مثل الطول لا يمكن أن يكون مضبوطا تماما ، ولكن من الممكن الوصول الى أقرب تقدير للقيمة الحقيقية للشيء المقاس وذلك بتقرير (أو تحديد) أحسن قياس يمكن أن تصل اليه الادارة المستخدمة ، مع أدراكنا أن القيمة الحقيقية قد تتأرجح

لمدى معين على جانبي هذا التقدير .

وهناك نقطتان جديرتان بالاهتمام :

الاولى :

أنه من الممكن انشاء مسطرة من هذه المجموعة من العصي اذا كانت مدرجة بدقة (أي أن هذه المجموعة تكون فيما بينها أداة قياس واحدة) . ولا يفوتنا بالطبع أن العصى في مثالنا قد درجت بواسطة مسطرة موجودة ، ولكن المهم هنا هو مفهوم التوصل من مجموعة من العصى الى كل موحد مثل المسطرة .

ثانيا :

أنه مع تحديدنا للعصا التي يتخطاها الشخص ٥٠٪ من المرات ، فأن العصى الاخرى من الممكن أن تعطينا صورة عن مدى فاعلية العصى في تحديد طول شخص ما . فأذا كانت أطوال العصى في هذه المجموعة تتغير بمقدار ربع بوصة فأنها تكون أدق كثيرا وأكثر حساسية في تقدير طول الشخص مما لو كانت أطوال العصى تتغير بمقدار بوصة كاملة .(١٧)

أن ترجمة تلك المبادئ من القياس الفيزيقي الى القياس السلوكي أسهل مما نتصور . أن العصى في المجموعة تصبح هي البنود في الاختبار ، حيث يوفق أو يفشل الفرد في الاداء على كل بند منها وحيث يتكون الاختبار (المسطرة) من هذه البنود مع بعضها . وكلما زادت حساسية الاختبار (أي دقة تدرج البنود) كلما كان الاختبار أكثر دقة كأداة للقياس .

نعود ثانية لنجمل نموذج راش .

١ ـ عندما يجابه فرد بندا فهناك نتيجة واحدة يمكن تسجيلها (نجح/ فشل).وهذه النتيجة تعتمد على :

أ) بعد الفرد : وهو ثابت لكل البنود التي يحلها هذا الفرد .

ب) بعد البند : وهو ثابت لكل الافراد الذين يجيبون على هذا البند .

٢ ـ أن بعد الفرد يقيس نفس ما يقيسه بعد البند ويعبر عنه على المقياس نفسه .

٣ ـ إن جميع البنود لاختبار ما تثير استجابات لدى الافراد على نفس السمة (ينبغي أن تقيس جميع البنود نفس الصفة) . (١٨)

لقد أرسى هذا النموذج ببساطة الشروط التي لو توافرت لامكن التوصل الى مقياس موضوعي للسلوك .

والان كيف يمكن تقدير بعدي نموذج راش ، بعد الفرد وبعد البند ؟؟؟

بعدا النموذج (وتقديرهما)

تقف بساطة نموذج راش عند المستوى الوصفي ، وذلك كما هو الحال في أغلب النماذج الرياضية ومع أن تحديد قيم بعدي النموذج (بعد الفرد وبعد البند) هي مهمة الحاسب الآلي ، فأن هذا لا يمنع من مناقشة الفكرة بصورة عامة .

بناء على النموذج ، يتصل قياس بعد البند مباشرة بنسبة عدد الافراد من عينة ما ، الذين لم يوفقوا في الاجابة على البند أجابة صحيحة ، وأن قياس بعد الفرد يتصل مباشرة بدرجة الفرد الكلية على الاختبار . وعلى هذا يمكننا أن نستخدم اصطلاحي (صعوبة البند)(Item Difficulty) ، و (أداء الفرد) بدلا من الاصطلاحين السابقين ولتوضيح السبب في أستخدام اصطلاح صعوبة البند بدلا من سهولة البند ـ كما هو المألوف . نرجع الى مثالنا السابق (العصى وطول الفرد) . أن طول العصى يزداد في نفس الاتجاه مع الشيء المقاس (طول الفرد) وليس من المألوف أن يحدد طول الفرد بعمق العصا وانما بطولها ، لذا يستخدم بعد صعوبة البند ، ما دام مستوى الصعوبة ودرجة التحصيل الاعلى يسيران في نفس الاتجاه . (١٩)

خطوات تقدير بعدي النموذج :

أولا : تصنيف الافراد الى مجموعات تضم كل واحدة منها هؤلاء الافراد الحاصلين على نفس الدرجة الكلية على الاختبار ، دون النظر الى أجاباتهم على كل بند ودون النظر الى مستويات العمر أو الفرق الدراسية .

ثانيا : البحث عما اذا كان هناك أي فرد قد حصل على درجة الصفر أو على الدرجة النهائية وعما اذا كان هناك بند أخطأ جميع الافراد في الاجابة عنه أو وفقوا جميعا في الاجابة عنه .

ولتوضيح ذلك نعود مرة اخرى الى مثال العصا وطول الفرد فأذا تخطى الفرد جميع العصى (أو لم يتخط أي عصا) فأننا لا نستطيع أن نعرف سوى أنه أطول أو أقصر من مدى هذه العصى ، أي أن هذا الشخص غير مناسب لهذه المجموعة من العصى ، ومن الممكن تقدير طوله بمجموعة أخرى من العصى . وعلى العكس اذا كان هناك عصاة قد تخطت جميع الافراد أو لم تتخط أي فرد منهم فأنها تكون عندئذ زائدة عن المجموعة وقد تناسب مجموعة أخرى من العصى .

وعلى هذا فينبغي حذف كل الافراد الذين يحصلون على الدرجة صفر أو الدرجة النهائية من التحليل حيث يكونون فوق أو تحت مدى الاختبار . كما أن جميع البنود التي يخفق في الاجابة عليها كل الافراد أو يجيبون عليها أجابة صحيحة فأنها تحذف أيضا من التحليل حيث تعتبر صعبة جدا أو سهلة جدا . (٢٠)

بعد ذلك يبدأ في قياس بعدي النموذج . وستكون المناقشة هنا فقط في الحالة التي تكون ـ الاجابة فيها ثنائية (صواب/ خطأ) ، علما بأن هذا النموذج يمكن ان يمتد الى مستويات اخرى من الاستجابة . من الفروض الرئيسية للنموذج وفي حالة ثنائية الاجابة يظهر تدريج نسبة مشترك بين كل من صعوبات البنود وأداء الافراد حيث من الممكن التعبير من مميز النجاح (Odds of Success) للفرد(S) على البند(i) وذلك بمصطلحات أداء الفرد(Zs) وصعوبة البند(Ei) هكذا .

$$O_{si} = \frac{Z_s}{E_i}$$

كما يمكن ان يعبر عن نجاح الفرد مرة اخرى كاحتمال النجاح مقسوما على أحتمال الفشل هكذا .

$$O_{si} = \frac{P_{si}}{Q_{si}}$$

$$\therefore \frac{P_{si}}{Q_{si}} = \frac{Z_s}{E_i} \qquad \therefore \frac{P_{si}}{Q_{si}} = \frac{Z_s / E_i}{1}$$

$$\therefore \frac{P_{si}}{P_{si} + Q_{si}} = \frac{Z_s / E_i}{1 + Z_s / E_i}$$

وعلى هذا فأن احتمال نجاح الفرد(s) على البند(i) يمكن أن يعبر عنه هكذا .

$$\therefore P_{si} = \frac{Z_s / E_i}{1 + Z_s / E_i}$$

أما في التطبيق العملي فأن الصورة اللوغاريتمية لكل من بعد الاداء(B v) وبعد الصعوبة (γi)تكون اكثر ملاءمة . عندئذ يمكن التعبير عن أحتمال نجاح الفرد(v) على البند(i) أي عندما تقدر استجابته(Xvi) بالدرجة هكذا .

$$Pr(X_{vi} = 1 \mid B_v, \gamma_i) = \frac{\exp(B_v - \gamma_i)}{1 + \exp(B_v - \gamma_i)} \qquad (١)$$

اما احتمال فشل الفرد(V) على البند(i) اي عندما تقدر استجابته(XvI) بالدرجة صفر فيمكن التعبير عنه هكذا .

$$pr(X_{vi} = 0 \mid B_v , \gamma_i) = 1 - \frac{\exp(B_v - \gamma_i)}{1 + \exp(B_v - \gamma_i)} \qquad (٢)$$

$$pr(X_{vi} = 0 \mid B_v , \gamma_i) = \frac{1}{1 + \exp(B_v - \gamma_i)}$$

من المعادلتين ١ ، ٢ تكون الصورة هي .

$$pr(X_{vi} \mid B_v , \gamma_i) = \frac{\exp(X_{vi}(B_v - \gamma_i))}{1 + \exp(B_v - \gamma_i)} \qquad (٣)$$

وفي الحالة الخاصة التي تتساوى فيها صعوبة البند مع مستوى أداء الفرد أي عند $B_v = \gamma_i$

فأن $B_v - \gamma_i = 0$

$$pr(X_{vi} = 1 \mid B_v , \gamma_i) = \frac{1}{1 + 1} = 0.5$$

اي احتمال نجاح ٥٠٪ (راجع مثال طول الفرد والعصى)

وعندماتوضع نتائج الاستجابات(X_{vi}) لمجابهة الفرد/ بند (بعضها واحد صحيح وبعضها صفر) ، تجمع الاعمدة وكذلك الصفوف لتعطي في نهايتها الدرجة الكلية لكل فرد كما تبين مجموع الافراد الذين أجابوا أجابة صحيحة عن كل بند .

X_{11}	X_{21}	X_{31}	X_{v1}	X_{N1}	$X_{.1}$
X_{12}	X_{22}	X_{32}	X_{v2}	X_{N2}	$X_{.2}$
X_{13}	X_{23}	X_{33}	X_{v3}	X_{N3}	$X_{.3}$
⋮	⋮	⋮	⋮	⋮	⋮
X_{1i}	X_{2i}	X_{3i}		X_{Ni}	⋮
⋮	⋮	⋮	⋮	⋮	⋮
X_{1K}	X_{2K}	X_{3K}	X_{VK}	X_{NK}	$X_{.K}$
$X_{1.}$	$X_{2.}$	$X_{3.}$	$X_{V.}$	$X_{N.}$	$X_{..}$

وعلى هذا فالدرجة الكلية للفرد(v) على الاختبار كله (جميع البنود) هي :

$$X_v = \sum_{i=1} X_{vi} \quad (٤)$$

كما أن مجموع الافراد الذين أجابوا أجابات صحيحة على البند(i) هو :

$$X_{.i} = \sum_{v=1}^{N} X_{vi} \quad (٥)$$

بأستخدام نظرية الترجيح الاكبر (Maximum Likelihood) نصل الى المعادلتين .

$$r = \sum_{r=1}^{k} pr \quad (X_{ri}) \quad (٦)$$

حيث(r) هي بالدرجة الكلية لكل شريحة(Total Score group)

$$S_i = \sum_{r=1}^{k-1} n_r \; pr \quad (X_{ri}) \quad (٧)$$

حيث(Si) هو عدد الافراد الذين أجابوا على البند(i) أجابة صحيحة(N r) هو عدد الافراد الذين حصلوا على الدرجة الكلية(r) .

وتحدد قيم(إδ) ، (B v) لكل بند ولكل فرد وهي القيم التي تحقق المعادلتين السابقتين آنيا .

عندئذ يكون هناك(K) من صعوبات البند (لكل بند δ من أبعاد الصعوبة) وعدد— (K 1) من أبعاد الاداء (B لكل درجة كلية محتملة) من 1 الى 1- k حيث لا يمكن تدريج الدرجتين الكليتين المتطرفتين صفر ، k وعلى هذا تحدد العلاقة بين الدرجات الكلية المحتملة وتقديرات أداء الفرد . حيث يمكن رصد كل درجة كلية محتملة على الاختبار في جدول مع مقابلها من مستوى اداء الفرد الذي يحصل على هذه الدرجة كما يمكن أن تحدد هذه العلاقة التقييسية(relation ship

Calibration)على شكل منحنى ويلاحظ هنا أن النموذج لا يفرق بين مستوى اداء شخصين لهما نفس الدرجة الكلية على الاختبار .

كما يمكن أيضا تكوين جدول لمستويات الصعوبة المقابلة لكل بند . وقد أطلق الاصطلاح «لوجيت» على وحدة القياس المشتركة لكل من تدريج أبعاد الاداء وأبعاد الصعوبة .

وينبغي الاشارة الى ان نمط الاجابة صواب/ خطأ هو أكثر الانماط مناسبة لاستخدام تكنيك الترجيح الاكبر(Maximum Likelihood) لتقدير الابعاد اللوغارتيمية لكل من صعوبة البنود وأداء الافراد آنيا وقد أدت طبيعة النموذج الاحتمالية الى أستخدام اللفظ تقدير(estimation) بدلا من اللفظ(evaluation) حيث أنه ليس هناك قياس دقيق بالمعنى المقصود (٢١) .

أخطاء القياس :

تحسب أخطاء القياس (الخطأ المعياري) لكل من صعوبة البنود وقدرة الافراد بواسطة تكتيك الترجيح الاكبر . وعندما يحدد الخطأ المعياري لصعوبة كل بند من بنود الاختبار ولأداء كل فرد على هذا الاختبار ، يمكن عندئذ تحديد الى أي مدى تتأرجح القيمة الحقيقية لصعوبة كل بند ولأداء كل فرد .

هل تتحقق الموضوعية :

تتضمن متطلبات الموضوعية في القياس أن يكون تدريج البند مستقلا عن الفرد(Person free) كما يكون تقدير أداء الفرد مستقلا عن البند(item free) حيث تحسب قدرة الفرد أو أدائه بحيث تكون ثابتة لفرد معين بصرف النظر عن أي من البنود التي يؤ ديها هذا الفرد ، كما تحسب صعوبة البند لتكون ثابتة لبند معين بصرف النظر عن الفرد الذي أداها وحين تحدد هذه القيم هل نصل الى ما يصبو اليه النموذج ؟

هل صعوبة البند مستقلة عن الفرد ؟ :

أن الاجابة عن هذا السؤ ال تتطلب مجموعتين من الافراد الذين يؤ دون المجموعة نفسها من البنود (الاختبار) . فأذا كانت قيم صعوبة البند الناتجة في كل مجموعة متكافئة أحصائيا (مع الاخذ في الاعتبار الخطأ المعياري لهذه الابعاد) ، عندئذ يكون الاحتمال الاكبر هو أن صعوبات البند مستقله عن الفرد . ومن الممكن تطبيق الاختبار الواحد لمجموعة واحدة من الافراد في جلسة واحدة ثم تجزئة هذه المجموعة الى مجموعتين حيث يمكن أستخدام وسيط التوزيع لدرجات الافراد الكلية على الاختبار ، فيكون في الغالب نصف الافراد واقعا في كل مجموعة فأذا حسبت صعوبات البنود من كل مجموعة من الافراد على حدة ، وكانت تلك الصعوبات متكافئة أحصائيا لكل من المجموعتين فهذا يعني أن صعوبات البند لهذا الاختبار مستقلة عن الفرد . (٢٢)

هل مقاييس الفرد مستقلة عن البند ؟ :

بطريقة مشابهة لتلك التي أتبعت للاجابة عن السؤ ال السابق يكون المطلوب هو مجموعتين من البنود (أختبارين فرعيين) تقيهان صفة معينة لدى مجموعة من الافراد . حيث تضم أحدى المجموعتين أصعب البنود وتضم الاخرى أسهلها . فاذا كان تحصيل كل فرد من هؤ لاء الافراد الذين قيس تحصيلهم بواسطة كل مجموعة من هذه البنود على حدة متكافئا أحصائيا فان هذا يعني أن مقاييس الفرد مستقلة عن البند .

ان تقسيم البيانات(data) الى قسمين (سواء الافراد أو البنود) ثم قياس مدى التكافؤ (التكافؤ الاحصائي الذي يأخذ في الاعتبار أخطاء القياس) وذلك للابعاد المقاسة (سواء كانت مقاييس صعوبات البند أو اداء الفرد) ، هو في الحقيقة أختبار حقيقي للملاءمة(fit) فاذا كانت البيانات ملائمة للنموذج فإن كل شيء يسير بصورة مرضية ، واذا لم يكن فينبغي مناقشته . (٢٣)

جراءات اختيار البنود :

قد يرجع النقص في ملاءمة البيانات موضوع الدراسة للنموذج الى مصدرين ـ سوء ملاءمة البنود وسوء ملاءمة الافراد . ولما كان من الممكن اختيار الافراد الملائمين لاختبار ما خاصة أن الاختبار الجيد ينبغي أن يغطي مدى كيفي وكمي واسع من القدرة دون أن يؤ ثر ذلك تأثيرا كبيرا على خواص التدريج ، لذا ينبغي التركيز على سوء ملاءمة البنود . وتعتبر ملاءمة البند للنموذج سيئة اذا كانت صعوبته بالنسبة لباقي البنود غير ثابتة عند مختلف مستويات أداء الافراد ، كأن يكون سهلا بالنسبة للافراد الصغار أو الأقل قدرة في حين يكون هو نفسه بندا صعبا بالنسبة للافراد الكبار أو العكس .

ان الهدف من تكوين أختبار جيد كما تحدده مصطلحات نموذج راش هو أن تكون لكل البنود القدرة نفسها على التمييز ، أو بعبارة اخرى أن تكون منحنياتها المميزة متوازية ، حيث المنحنى المميز للبند(ICC) (Item charactristic curve) هو رسم بياني لاحتمالات النجاح على البند وذلك للافراد ذوي المستويات المختلفة من الاداء . ويشير الى أن أحسن البنود ملاءمة للنموذج ليست بالضرورة هي المتقبلة من حيث شكل منحنياتها المميزة(ICC) . لهذا ولمثل هذه الحالات نكون مضطرين لاختيار بنود معتمدين على شكل منحنياتها المميزة ، ثم بعد ذلك تدخل هذه المجموعة المنتقاة من البنود في برنامج الكمبيوتر حيث تكون خطوات أختيار البنود تلقائية (أوتوماتيكية) . وكما أشار(Murray) في تقريره عن أختبار الذكاء البريطاني القومي ، الى أن البنود التي يتكرر رفضها على أساس ضعف ملاءمتها لنموذج راش هي نفس تلك البنود التي يمكن أن ترفض على أساس ضعفها بالنسبة للمنطق السيكلوجي ، أو على أساس عدم تجانسها مع غيرها من البنود من حيث المضمون ، وهذا يؤ كد الثقة في صدق عامل القياس النفسي في اجراءات

اختيار البنود . (٢٤)

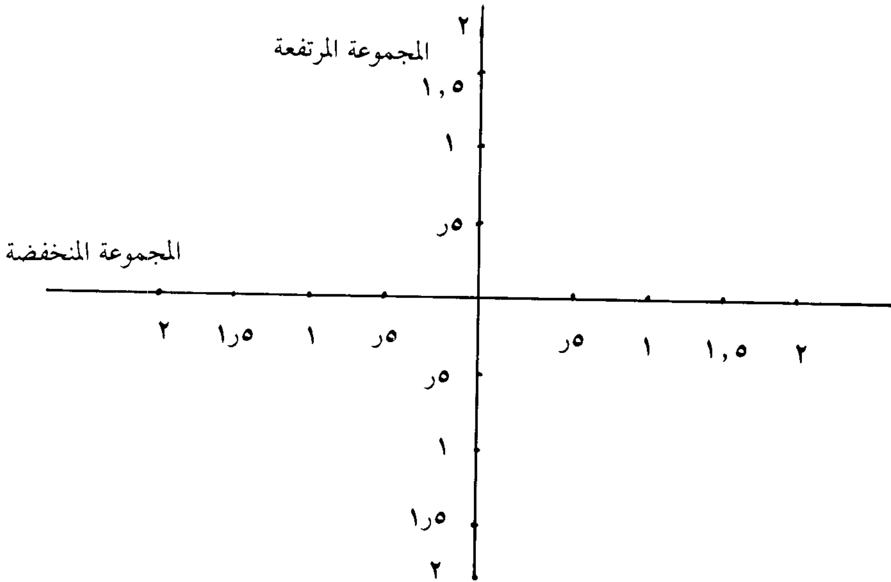
مثال لتطبيق نموذج راش على احد الاختبارات :

قد يكون من المناسب عرض تطبيق للنموذج على أحد الاختبارات كي نرى كيف يمكن الوصول الى مجموعة من البنود (أختبار) تحقق فروض النموذج وما تهدف اليه من موضوعية في القياس .

لنأخذ أحد الاختبارات التي عرضها كل من (Willmot ,A . and Fowles ,D .,) وهو اختبار في الفهم اللغوي (English comprehension) وهو من (Model C .S.E . Exam in ation) لسنة ١٩٧٢ بانجلترا ، وهو أختبار مكون من ٤٠ بندا . أختيرت ١٠٠٠ ورقة أجابة من أوراق الطلبة الذين أدوا هذا الاختبار بحيث يكون هناك عددٌ متساوٍ من الافراد الحاصلين على كل درجة كلية يحتمل الحصول عليها (حيث الدرجة الكلية هي عدد البنود التي يجيب عنها الفرد اجابة صحيحة ، نصفت أوراق الاجابة بناء على الدرجة الكلية ، بحيث أصبح هناك مجموعتان منفصلتان ، واحدة تضم الحاصلين على الدرجة ١٩ فأقل (المجموعة المنخفضة) والاخرى تضم الحاصلين على الدرجة ٢٠ فأكثر (المجموعة المرتفعة) . عندما حسبت معاملات السهولة بالطريقة التقليدية ، تبين اختلاف واضح لهذه المعاملات لدى ـ المجموعتين . حيث كانت سهولة البند أكبر دائما لدى المجموعة المرتفعة عنه لدى المجموعة المنخفضة ، فقد كانت سهولة البند الاول على سبيل المثال مساوية ٧٦, ٠ لدى المجموعة المرتفعة في حين كان ٤٢, لدى المجموعة المنخفضة ، مما يدل بوضوح على مدى تأثير مستوى العينة على سهولة البند ، كما يبدو من تحليل البنود بالطريقة التقليدية . ترى هل يمكن الوصول الى صعوبات البنود بحيث تكون مستقلة عن أثر العينة ؟

عندما تبدأ مواجهة الفرد/ بند تكون الاجابة أما صوابا وأما خطأ وعندما ترصد نتائج أستجابة كل فرد لكل بند من البنود يمكن التوصل الى تقديرات لصعوبة كل بند . وتحسب هذه التقديرات من كل مجموعة من المجموعتين على حدة (المرتفعة والمنخفضة) وترصد في جدول حيث لوحظ أنها كانت تتأرجح بين + ٢، - ٢ لوجيت يقارن بعد ذلك تقديري الصعوبة لكل بند الناتجين من كل مجموعة مع الاخذ في الاعتبار قيمة الخطأ المعياري .

كما يمكن فى رسم بياني تحديد مواقع النقاط التي تمثل تقديرات صعوبة كل بند كما تقدر من كل من المجموعتين المرتفعة والمنخفضة .



وتقع البنود التي يتكافأ فيها بعدا صعوبتها المحسوبان من كل من المجموعتين على الخط المار بنقطة الاصل والذي يميل ٤٥° على كل من المحورين وهي البنود المستقلة عن أثر العينة .

وقد أوضح كل من الجدول والرسم أن تلك البنود التي يتكافأ فيها بعدا الصعوبة تبلغ ٢٣ بندا من ٤٠ بندا هم أصل الاختبار ، في حين أن ١٧ بندا يختلف فيها بعدا الصعوبة التي تتأثر بمستوى المجموعة التي أجرى عليها الاختبار . وقد عملت أجراءات اكثر شدة ، وذلك بأخذ نصف الافراد الاقل مستوى وذلك من المجموعة المنخفضة ونصف الافراد الاعلى مستوى وذلك من المجموعة المرتفعة . وأعيدت جميع الاجراءات السابقة لتحديد أبعاد صعوبات بنود الاختبار ، وذلك على كل من المجموعتين الجديدتين من الايراد . وكانت النتيجة مطابقة لما سبق ما عدا بندا واحدا . وعلى هذا أصبح هناك ٢٢ بندا متشابهة أحصائيا من حيث أبعاد صعوباتها بالنسبة لكل مجموعة ، أي تعتبر بنودًا غير متأثرة بمستوى العينة ، أي تلائم النموذج ، في حين كان هناك ١٨ بندا غير متشابه ، أي متأثرة بمستوى العينة ، ومن ثم فهي لا تلائم النموذج . وقد تبين بعد الفحص الدقيق أن بعض البنود غير الملائمة للنموذج كانت غير واضحة الصياغة أو أنها تقيس شيئا آخر غير الفهم اللغوي وقد كان هناك ١٥ بندا من هذه البنود الثمانية عشر غير ملائمة لاسباب متقاربة .

أما الـ ٢٢ بندا الملائمة للنموذج فقد كانت تقيس نفس الصفة وهي الفهم اللغوي (English Comprchension) . وأعيدت الاجراءات مرة أخرى بعد حذف الـ ١٨ بندا وأعتبار الاختبار مكونا من ٢٢ بندا بدلا من أربعين ، فوجد أن صعوبات البنود الناتجة مشابهة جدا ، لصعوباتها عندما كانت ضمن الاختبار المكون من ٤٠ بندا ، وان هذه البنود الاثنين والعشرين ما زالت ملائمة لمتطلبات النموذج .

العلاقة التقيسية بين مجموعة الدرجة الكلية وتقدير أداء الفرد :

ان حساب صعوبات البند من نتائج مجابهة الفرد/بند يعطى في الوقت نفسه تقديرا لمستوى الفرد (حيث تحل المعادلتين ٦ ، ٧ آنيا) . وعلى هذا فلكل من المجموعتين ـ المرتفعة والمنخفضة ـ يمكن استخلاص العلاقة بين مجموعات الدرجة الكلية(Total score group) وتقدير أداء الفرد الحاصل على هذه الدرجة الكلية .

وتعرف هذه العلاقة بالعلاقة التقيسية(Calibration relationship) ويمكن رصد هذه العلاقة في جدول مع بيان الخطأ المعياري لابعاد أداء الفرد (حيث الخطأ المعياري هنا دالة لعدد بنود الاختبار ، وأيضا لنسبة الافراد في مجموعة درجة كلية معينة الذين يحلون صوابا كل بند من البنود) .

وقد تبين أن مستوى الاداء يتراوح بين حوالي + ٤ ، - ٤ «لوجيت» كما يمكن تصوير هذه العلاقة برسمها في منحنى . وقد عمل المنحنيان العياريان أو التقيسيان(Calibration Curves) الناتجان من كل من مجموعتي الافراد (المرتفعة والمنخفضة) حيث تبين انهما متطابقان مع الاخذ في الاعتبار الاخطاء المعيارية .

هل تقدير الفرد مستقل عن البند :

ينبغي التحقق من الشرط الآخر للموضوعية ، الذي يحقق النموذج ، وهو أن تقدير مستوى أداء الفرد يكون مستقلا عن البند الذي يجيب عليه . وللتحقيق من ذلك قسم الاختبار (**٢٢ بندا**) الى اختبارين فرعيين ، يشمل احدهما أسهل البنود (حسب تقدير أبعاد الصعوبة) ، ويشمل الآخر أصعب البنود وقد عملت جميع إجراءات تحديد أبعاد صعوبة البنود وأبعاد أداء الفرد مرتين وذلك بأستخدام كل اختبار فرعي على حدة وقد أستخرجت العلاقة بين مجموعة الدرجة الكلية ومستوى أداء الفرد لكل اختبار فرعي على حدة ، بالطبع لا يتوقع أنه اذا تساوت الدرجة الكلية على كل اختبار فرعي أن يتساوى مستوى أداء الفرد ، لأن أحد الاختبارين سهل والآخر صعب ، ولكن ما يتوقع هو أن للفرد الواحد درجتين كليتين مختلفتين على كل أختبار ، ولكن كلا منهما يقابل نفس مستوى الاداء للفرد عندئذ نستطيع القول بأن اداء الفرد مستقل عن مستوى مجموعة البنود التي يؤ ديها .

بعض مميزات أخرى للنموذج :

١ ـ تعديل مدى التدريج :

ان الاجراءات المتبعة في هذا النموذج لتقدير الأبعاد جعلت المجموع الكلي لابعاد صعوبات البنود مساويا للصفر .

ولما كانت أبعاد كل من صعوبة البندوأداء الفرد قد قدرت على مقياس من الوحدات المتساوية (Interval scale) كما سبقت الاشارة ـ فمن الممكن تعيين نقطة الصفر أصطلاحيا ، كما يمكن تحويل كلا من أبعاد الصعوبات وأبعاد القدرة خطيا ، لتكون في مدى اكثر مناسبة ، كان تكون من صفر الى ١٠٠ ، وهذا يشبه تحويل تدرج الحرارة الفهرنهيتي الى المئوي والعكس .

٢ ـ ضم البنود وتكوين بنوك لها :

كما ذكرنا سابقا فان المجموع الكلي لصعوبات مجموعة البنود (الاختبار) مساوية للصفر . فاذا كانت هناك مجموعة أخرى من البنود (ولتكن مجموعة أصعب) درجت بالطريقة نفسها على العينة ذاتها أو عينة أخرى من الافراد فان متوسط صعوبة هذه المجموعة من البنود يكون صفرا ايضا . اذن كيف يمكن أن توضع البنود على التدريج نفسه من حيث الصعوبة ؟

وتبدو أهمية هذا السؤ ال بوضوح عندما يكون مدى الاختبارات الفرعية أكبر من أن تدرج كلها في أن واحد على عينه الافراد نفسها أو عندما يراد تكوين بنك للبنود حيث يكون من المرغوب فيه أضافة بنود جديدة دون الاضطرار الى اعادة تدريج بنود البنك .

ان خواص نموذج راش في تدريج صعوبات البنود يتيح الضم بين مجموعتين من البنود ، فبالرغم من أن نقطتي الصفر لكل من المجموعتين اصطلاحيتان، فان حجم الوحدة لكل من التدرجين واحدٌ . وبتعبير آخر فان مقدار ازاحة واحد «لوجيت» على كل من التدريجين لها المعنى نفسه بأصطلاحات مميزة النجاح وهذا يعني انه لوضع كل من المجموعتين على تدريج واحد ، يستدعي فقط اضافة أو طرح مقدار ثابت على كل من صعوبات البند لأحدى هاتين المجموعتين . ويمكن الحصول على قيمة هذا الثابت من مقارنة صعوبة البنود المشتركة بين كل من المجموعتين . وعلى هذا فان أستخدام عينتين مختلفتين من الافراد ومجموعتين متداخلتين من البنود يمكننا من تدريج أختبار يغطي مدى كبيرا من الصعوبة أكثر مما يمكن أن يدرج بعملية واحدة على عينه واحدة .

وقد تتكرر عملية الضم هذه عدة مرات لتغطية مدى أوسع من الصعوبة ولكن كلما زادت عمليات الضم زادت الاخطاء وتراكمت ، ومن الممكن التقليل من أخطاء الضم وذلك بواسطة عمليات تسمى Jack Knifing (٢٦) .

٣ ـ الصورة المختصرة للاختبار وحبك الاختبارات :

(Short and Individualised tests :)
(٢٧)

ان الصورة المختصرة للاختبار تعني أستخدام كل ثان أو ثالث بند من بنود الاختبار ، مع التسليم بأن كل بند قد وضع بناء على مستوى صعوبته ـ في مكانه الصحيح في الاختبار الاصلي ،

وأن متوسط صعوبة البنود للاختبار في صورته المختصرة هو نفسه متوسط صعوبة البنود للاختبار الاصلي كله . ومن هنا فأنه يمكن بكل بساطة أن تضرب الدرجة الكلية التي حصل عليها الفرد على الاختبار في صورته المختصرة × ٢ أو × ٣ ثم مراجعة الجدول الذي يوضح العلاقة بين الدرجة الكلية ومستوى أداء الفرد وذلك للاختبار الكلي .

وهناك اجراءات صادقة لحبك الاختبار للفرد ، حيث يبدأ الفرد عند نقطة يظن أنها مناسبة ، فاذا فشل الفرد في هذا البند يحاول في آخر أسهل ، فاذا نجح فيه ينتقل الى الاصعب وبأعادة هذه العملية عدة مرات نحصل على بيانات لأداء هذا الفرد على مجموعة من البنود تتمركز تقريباً حول قدرته ويمكن أن تحسب قدرة الفرد أو مستوى أدائه بالرجوع الى الجدول المتدرج للأختبار ذي الطول المعين (الجدول الذي يوضح العلاقة بين الدرجة الكلية للفرد على البنود ومستوى قدرته أو أدائه) .

كيفية الاستفادة من تطبيق نموذج راش في البيئة المصرية :

سنحاول فيما يلي أن نناقش كيفية الاستفادة من تطبيق نموذج راش في كل من مجال القياس التربوي ومجال القياس النفسي في المجتمع المصري .

أولا : في مجال القياس التربوي :

أ ـ قياس التحصيل الجامعي :

ازداد عدد الجامعات الاقليمية في الآونة الاخيرة زيادة كبيرة ، وانشئت الكثير من الكليات المناظرة لغيرها بالجامعات الاخرى . وقد أستقطبت هذه الكليات الطلبة الممتازين بالاقاليم ، كما حظيت تلك الجامعات الاقليمية بالميزانيات الانشائية والنصيب الاكبر من البعثات . . . الى آخره من مظاهر الاهتمام . مع هذا هناك من يعتقد بأن خريجي الجامعات الجديدة ليسوا في مستوى خريجي الجامعات الاولى ، بل هناك من ينادي بألا ينضم خريجو بعض الكليات الى النقابات المهنية وذلك بهدف منعهم من ممارسة المهنة (كما حدث في نقابة الأطباء) . ولسنا هنا بالطبع بصدد مناقشة هذا الرأي ولكنا نورده لما يحمل في طياته من المقارنة بين خريجي الكليات المتناظرة في الجامعات المختلفة من محافظات قد تتباين مجتمعاتها .

هنا تبرز أهمية تطبيق نموذج راش في انشاء المقاييس الموضوعية التي تقدر مستوى أداء الفرد بوحدة قياس محددة (لوجيت) . فمن الممكن مثلا انشاء بنك للبنود خاص بكل مقرر من المقررات الجامعية .

فيمكن مثلا انشاء بنك بنود لمقرر المحاسبة ، يغطي كل ما يدرس في هذا المقرر بالجامعات المصرية ، وذلك بتطبيق نموذج راش مما يتيح معرفة مستوى صعوبة كل بند من هذه البنود ـ حيث تقع على مقياس مدرج واحد ـ كما يتيح معرفة مستوى أداء الفرد أو تحصيله اذا حصل على درجة

كلية معينة على مجموعة معينة من هذه البنود . ويستطيع بالطبع كل أستاذ أن يختار من هذا البنك مجموعة البنود التي تحقق هدفه من تدريس هذا المقرر . ولقد ناقشنا من قبل الاجراءات التي تتبع لمعرفة مستوى أداء أو تحصيل الفرد من معرفة درجته الكلية مهما اختلف عدد البنود التي يؤ ديها (الصورة المختصرة للاختبار) .

هنا نستطيع مقارنة أداء الافراد بعضهم ببعض مهما اختلفت كلياتهم أو جامعاتهم ، ما داموا قد درسوا مقرر المحاسبة أوجزءاً منه ، وما دام بنك البنود يغطي ما درسه هذا الفرد وهذا ما يعجز عن تقديمه الاختبار بصورته التقليدية جماعية ـ المرجع ، حتى لو كان هذا الاختبار التقليدي له صورتان

ولا يعوق هذا النموذج الاستخدام عندما يكون الهدف هو مقارنة أداء الفرد بمحك أو مستوى معين (كذا لوجيت) نرى أنه ينبغي أن يصل اليه الفرد لكي يعتبر ناجحا في دراسة هذا المقرر ، أو بمقارنته بمحك أو مستوى آخر نرى انه ينبغي أن يصل اليه الفرد لكي يستكمل دراسته العليا في مقرر المحاسبة . ولكن قد تثور بعض التساؤلات عند أستخدامها كمحكية ـالمرجع حيث أحسن البنود في الاختبار هي تلك التي يحصل عليها الافراد على الصفر قبل دراستهم لمقرر المحاسبة ثم يحصلون على الدرجة النهائية بعد دراستهم لذلك المقرر في حين تحذف هذه البنود من الاختبار الذي يبنى بتطبيق نموذج راش .

ب ـ قياس التحصيل في التعليم العام :

من الممكن تطبيق نفس الفكرة السابقة وذلك على مستوى التعليم العام أيضا ، خاصة بعد الاعداد الكبيرة الهائلة من الطلبة والطالبات والتي يستحيل معها أجراء أمتحانات مركزة بالعاصمة ، والذي قد يؤدي الى عقد أمتحانات مستقلة بكل مديرية تعليمية . هذا مع ملاحظة تزويد البنك دائماً ببنود جديدة ، حيث يستعان ببعض البنود الموجودة بالبنك في عملية تدريج البنود الجديدة بحيث تكون معها على نفس المقياس المدرج .

وقد سبقت الاشارة الى كيفية ضم البنود الجديدة وتكوين بنوكها .

ثانيا : في مجال القياس العقلي :

يتكون المجتمع المصري من شرائح متباينة ، خاصة في المستوى التعليمي ، فمن الاميين الذين لا يقرأون ولا يكتبون الى ذوي المؤهلات العالية بل والعلماء والخبراء العالميين . ان عمل مقياس قومي للذكاء في مصر لهدف نتوق اليه ، ولكنه مستحيل التحقيق بالطرق الراهنة التقليدية في القياس النفسي ، حيث يستحيل أجراء نفس الاختبار على شرائح المجتمع المختلفة . وكم أخفق علماء النفس في مصر في محاولاتهم لعمل معايير لاختبارات الذكاء المقننة على بعض شرائح المجتمع وفي شرائح أخرى منه .

ولكن هناك أمكانية عمل أو أنشاء مقياس قومي للذكاء في مصر يمكن تطبيقه على جميع شرائح المجتمع المصري وذلك بتطبيق نموذج «راش» بعد تذليل مشاكل التطبيق التي سوف نتعرض لها في مجال نقد النموذج . فمن الممكن أنشاء عدة اختبارات متداخلة (أي تشترك في بعض بنودها) ، حتى يمكن ربطها بعضها ببعض وتدريجها على نفس المقياس المدرج حتى ننتهي الى مقياس واحد يغطي مدى واسعا من قدرة عقلية معينة ، ثم يكرر هذا بالنسبة الى قدرة عقلية ثانية وثالثة . . . . حتى يمكن أن نصل الى مقياس للذكاء يغطي مدىً واسعا من القدرات كما يغطي مدىً واسعا من المستوىً لكل قدرة من هذه القدرات .

مدىً الأستفادة بتطبيق نموذج راش في جامعة الكويت :

تتبع جامعة الكويت نظام المقررات وهي الاولى بين جامعات الخليج في تطبيق هذا النظام الذي يتيح للطالب الفرصة في اختيار المقررات المناسبة في اطار الخطة الموضوعية . فهناك متطلبات جامعية وهناك مقررات أجبارية للتخصص وهناك مقررات أختيارية في أطار هذا التخصص ، وهناك مقررات لها متطلبات من مقررات أخرى وكل هذا موضوع بحيث يحقق أهدافا عامة للثقافة الجامعية وأهدافا خاصة بكل كلية وكل قسم بما يحقق أهداف المجتمع وأهداف الطالب .

وفي هذاالنظام المرن يجد الطالب نفسه امام اختيارت عديدة للمقررات والمواعيد المناسبة والأساتذة القائمين على التدريس . وكثيرا ما نرى أقبال الطلبة علىً أستاذ معين أما لقدرته الفائقة كأستاذ واب ورائد وأما لأن تقديراته التي يعطيها على قدر من السخاء يتيح للطالب معدلات عالية . هنا نصل الى بيت القصيد فهناك مقرر معين أو مقررات متناظرة يقوم بالتدريس فيها أساتذة مختلفون ، عندئذ نجد فروقا في التقدير تختلف بأختلاف الأساتذة وطرق تدريسهم وأختلاف الامتحانات التي يقدمونها ثم اختلاف تقديراتهم لاجابات الطلبة على هذه الامتحانات . ولنأخذ مثالا لذلك مقرر (١٠١) في علم النفس حيث يقوم سبعة عشر أستاذا على تدريسه . هنا يتيح تطبيق نموذج راش أنشاء بنك للبنود المقدرة الصعوبة تغطي كل أهداف المقرر وتتناول جميع الخبرات وأشكالها المختلفة التي يتعرض لها كل هؤلاء الأساتذة . ولقياس تحصيل الطلبة يسحب كل أستاذ مجموعة البنود التي تحقق الأهداف التي تتفق مع الخبرات التي يدرسها مع الطلبة مكونا بذلك أختبارا يجيبون عليه ويقدر لكل طالب منهم الدرجة الكلية على هذا الأختبار ثم يرجع الى الجدول الذي يوضح العلاقة التقيسية بين الدرجة الكلية للطالب على الأختبار ومستوى أدائه مقدرا باللوجيت .

في هذا الغاء للفوارق في تقدير الطلبة بين هؤلاء الأساتذة المختلفين حيث يقدر لكل طالب مستوى أداء على هذه المادة تقديرا موضوعيا لا يتأثر بالبنود التي يجيب عليها الفرد ولا بالأفراد الذين يجيبون على الأختبار ، وبالطبع لا يتأثر بالأشخاص الذين يقومون بتصحيح الأختبار .

كما أن تطبيق نموذج راش على امتحان الثانوية العامة بالكويت ـ وتحديد مستوى معين تقرر على أساسه قواعد القبول بالجامعة بحيث من اللازم أن يكون مستوى أداء الطالب في مادة معينة

كذا لوجيت وفي مادة أخرى كذا لوجيت وهكذا ، وذلك للقبول بكلية معينة ـ يكون أكثر عدالة من مجرد الأعتماد على النسبة المئوية للمجموع الكلي للطالب في امتحان قد يختلف مستواه ويتغير من عام لآخر .

هل نهمل التراث العربي من القياس العقلي أم نستفيد منه ؟

فمن الممكن بالطبع أن تستفيد من التراث الموجود من المقاييس العقلية العربية عند انشاء مقاييس ملائمة لنموذج «راش» فاذا أخذنا على سبيل المثال انشاء مقياس للقدرة العددية يلائم نموذج «راش» فيمكن أن نتبع الخطوات الآتية :

١ ـ حصر جميع الأختبارات التي تقيس القدرة العددية بجميع مستوياتها وأشكالها .

٢ ـ نبدأ بأختبار منها وليكن (أ) ويجري الأختبار على مجموعة كبيرة متباينة من الأفراد يفترض أنهم على مدى واسع من القدرة العددية .

٣ ـ نطبق عليه أحراءات نموذج «راش» حتى نصل الى الصورة التي نطمئن الى أنها تلائم هذا النموذج .

٤ ـ يكرر نفس العمل بالنسبة الى الأختبارات التي يفترض أنها تقيس أيضا القدرة العددية . ولكن ينبغي ملاحظة أن تكون هناك بنود مشتركة بين الأختبارين حتى يمكن تدريج صعوبة بنود الأختبار على نفس المقياس المدرج للأختبار (أ) .

٥ ـ وهكذا يمكن تكرار نفس العمل حتى نستغرق ما عندنا من أختبارات تقيس القدرة العددية مع تعديلنا للبنود التي لا تلائم النموذج وأدخالها مرة أخرى في بنود الأختبارات .

وهذا يستدعي أيضا عمل بنود جديدة تملأ ما قد يكون من فجوات ومن هنا يتكون بنك لبنود القدرة العددية مدرجة جميعها على نفس المقياس المدرج للقدرة العددية . بحيث اذا طبقت مجموعة من البنود التي تلائم فردا معينا نستطيع تقدير مستوى قدرته العددية بمقدار (كذا لوجيت) مع أخذنا في الأعتبار مسألة عدد البنود التي يجريها الفرد .

نقد نموذج راش

بعد أستعراضنا للنموذج الذي ابتدعه «جورج راش» كي يحقق الاهداف المختلفة للقياس ولا يتعارض مع مسلماته المتباينة ، وذلك بمحاولته الوصول الى الموضوعية البحتة في القياس السلوكي مشتقا أفكاره من مشابهته الظواهر السلوكية بظواهر الحياة الفزيائية ، ينبغي أن نقوم الموقف ونضع هذا النموذج على كفة الميزان .

وعندما نحاول أن ننقد هذا النموذج فأننا سنتناول هذا من النواحي الاتية :

ـ مناقشة بعض مسلماته الاساسية .

ـ تقييمه من حيث مدى تحقيقه لبعض الاغراض التي وضع من أجلها .

ـ مناقشة بعض الصعوبات التي تكتنف تطبيقه .

١ ـ مناقشة بعض مسلماته الأساسية :

أ ـ ان الأفراد ذوو قدرة أحادية البعد :

هناك علامة أستفهام كبيرة على هذه المسلمة . فهل من الممكن أن يكون تحصيل الافراد وقدرتهم ذا بعد واحد مثل أطوالهم وأوزانهم ؟ فعندما ننتقل من المستوى الفزيائي الى المستوى البيولوجي ، ومن المستوى البيولوجي الى المستوى السيكلوجي ، فاننا لا نستطيع القول بأن هذه النقلة تمثل نقلات كمية فحسب بل هي أيضا نقلات كيفية . ومن ثم يكون تشبيه الاستجابات السلوكية التي تعتمد على اللغة وعلى المعنى (المستوى الثالث) ، بصفة كالطول مثلاً (المستوى الفيزيائي) تشبيها يعوزه الدقة ويتسم بالاختزالية ، أي بتسطيح المشكلات المعقدة وأختزالها اختزالا ـ قد يكون مخلا ـ الى جزئيات تنتمي الى مستويات أدنى . أن الطول صفة أحادية البعد أما أساليب السلوك فلها مضمون ثقافي حضاري يجعل الافراد يختلفون في تلك الاساليب تبعا لانتماءاتهم الجماعية ، مما يجعل فرض نموذج معين للاجابة صح/ خطأ ، مثيرا للمناقشة . وقد ظهرت من قبل الاختبارات المتحررة من الثقافة(Culture free) ولكن وجد أنه من غير الممكن أن ينسلخ الفرد عند أجابته عن مفردات مقياس ما ـ عن المؤثرات الثقافية التي نشأ فيها وترعرع . فظهرت بعدها المقاييس العادلة أو المتعادلة الثقافة(Culture fair)لعلها تحقق الهدف ، ولكن في الحقيقة ، لا يمكن لاختبار أن يكون عادلاً الا عند الاخذ بمسلمة معينة وهي أن يمر جميع الافراد بنفس الخبرات أي أن تكون حياة بني الانسان نمطا واحدا . من هنا نرى أن مسألة التحرر من أثر الثقافة وبالتالي أن يكون جميع الافراد على متصل من القدرة أحادي البعد مسألة قد يكون فيها شك كبير .

ان المناقشة السابقة تصح ، اذا سلمنا بأن الظاهرة السلوكية التي نقيسها هي ظاهرة سلوكية بسيطة ، فما بالنا اذا كانت الظواهر السلوكية مهما صغرت وبسطت ففيها قدر من التعميم والتعقيد .

ب ـ حذف الافراد الحاصلين على درجة الصفر والدرجة النهائية على أختبار ما :

عندما يجري الاختبار بغرض معايرته أو تقييسه بطريقة نموذج «راش» فينبغي حذف الافراد الحاصلين على درجة الصفر أو على الدرجة النهائية على هذا الاختبار حيث يعتبر هؤلاء الافراد خارج مدى الاختبار (درجة تجمد الماء درجة غليانه خارج مدى الترمومتر الطبي) . ولكن يحرمنا هذا الاجراء من أهم خصائص الاختبارات محكية ـ المرجع . حيث تعكس التغير أو النمو أو

الاكتساب ، حين يخفق الافراد في الاجابة عن جميع بنود الاختبار قبل البرنامج التربوي مثلا ثم ينجحون في الاجابة عن جميع بنود الاختبار بعد ذلك البرنامج .

جـ ـ حذف البنود التي لا يجيب عليها أحد أو التي يجيب عليها جميع الافراد :

بالمثل يستدعي معايره أو تقييس اختبار ما بتطبيق نموذج «راش» ، استبعاد تلك البنود التي يخفق جميع الافراد في الاجابة عنها أو تلك التي ينجح جميعهم في الاجابة عنها بحيث تعتبر هذه البنود، غير صالحة لقياس الافراد وهذا بدوره يحرمنا من أهم خصائص الاختبارات محكية ـ المرجع ، حيث تعتبر أحسن البنود هي تلك التي لا يستطيع جميع الافراد الاجابة عنها (معامل السهولة صفرا) قبل البرنامج التربوي مثلا ثم يستطيع جميع الافراد أن يجيبوا عن تلك البنود (معامل السهولة واحدا صحيحا) بعد ذلك بعد البرنامج ، أي تلك البنود التي تستطيع أن تعكس التغير أو النمو أو الاكتساب .

٢ ـ تقويم النموذج من حيث مدى تحقيقه لبعض الاغراض التي وضع من أجلها :

من الشروط التي يسعى الى تحقيها نموذج «راش» للوصول الى الموضوعية في القياس أن يكون تقدير صعوبة البند مستقلا عن عينه الافراد ، وأن يكون تقدير أداء الفرد مستقلا عن البند الذي يجيب عليه . وقد بدا من مناقشة الشرط الأول أنه ممكن التحقيق تماما أما الشرط الثاني فلم يكن تحقيقه بنفس هذه الصورة .

وقد أعطى مثالا على أختبار الطبيعة في أمتحان(G . C . E ., O -Level) حيث أستبعد ٢٠ بندا غير ملائمة للنموذج ، وأستبقى ٥٠ بندا أعتبرت ملائمة له . ثم قسمت البنود الى جزئين ، أحدهما يضم البنود السهلة والاخر يضم البنود الصعبة . ثم أجرى الاختباران الفرعيان على ٧٤٥ فردا ، فأصبح لكل واحد منهم درجتان ، درجة على كل اختبار فرعي . وحيث ان هاتين الدرجتين مستقلتان لذا فان تقدير تحصيل الافراد على كل واحد من الاختبارين الفرعين يكون مستقلا عن الاخر . وقد تبين ان ٩٤,٤٪ من هؤلاء الافراد كان تقدير تحصيلهم على كل اختبار فرعي **متساويا** إحصائياً، في حين ان ٥,٦٪ منهم كان تقديرهم على كل اختبار فرعي **مختلفاً** . أي أن اكثر من ٥٪ من الافراد لا يلائمون هذا النموذج الاحتمالي . هنا يشير(Willmot ,A . and Fowles) ، الى أن هذه النسبة (٥٪) تشبة مستوى الدلالة ومن هنا تكون هذه النتيجة منطقية متوقع . (٢٨)

ب ـ تقدير الاتجاهات والقيم :

كانت التطبيقات لنموذج «راش» منصبة أما على القدرات في المجال النفسي ، وأما على التحصيل في المجال التربوي . ولكن لم يكن هناك ما يشير الى تطبيق نموذج «راش» في مجال الاتجاهات أو القيم . ولعل التشبع الواضح للاتجاهات والقيم بثقافة المجتمع ،؛ كما يبدو ذلك

من مراحل تكوين الاتجاهات والقيم اثناء عملية التنشئة الاجتماعية ، وعدم القـدرة على تحـديد الصواب والخطأ جعل تطبيق هذا النموذج منصبا على تقدير القدرات والتحصيل .

مناقشة بعض الصعوبات التي تكتنف تطبيق النموذج :

أ ـ تلاشي الاختبار !!

عندما تجري معايرة أختبار ما أو تقييسه بواسطة نموذج «راش» فينبغي كما سبقت الاشارة اليه ، أن يحذف الفرد الذي يعتبر أداؤ ه خارج نطاق الاختبار مما قد يؤ دي الى أن يصبح احد بنود الاختبار خارج نطاق العينة ، وعندئذ نضطر الى حذف هذا البند . وعلى ذلك قد يصبح هناك فردا آخر خارج نطاق الاختبار وهكذا . فإذا كانت الدرجة النهائية هي الدرجة التي حصل عليها أحد الافراد فينبغي أن يحذف من التحليل ، فاذا حدث أن هذا الفرد كان الوحيد الذي أجاب على بند معين أجابة صحيحة ، فان حذفه يؤ دي الى عدم وجود اى فرد قد وفق في الاجابة على هذا البند فينبغي عندئذ حذف هذا البند . وقد يترتب على ذلك ان **يصبح هناك فردٌ** كان الخطأ الوحيد لديه هو هذا البند ، وعندما يحذف هذا البند يصبح هذا الفرد حاصلا على الدرجة النهائية وهنا يحذف هذا الفرد . وهكذا فاذا استمر هذا الامر بهذه الصورة فقد نصل الى أن نحذف بنود الاختبار كلها وينتهي أمرها ويتلاشى الاختبار ! (٢٩)

ب ـ الصعوبات العملية :

على الرغم مما ظهر من مناقشة نموذج «راش» فان النظرية التي يقوم عليها تبدو فعالة ولكن هناك الكثير من الصعوبات التي تكتنفها في مجـال التـطبيق العملي . ولهـذا بدأت مجموعـة من الابحاث بالمركز القومي للبحوث التربوية بأنجلترا وويلز(N . F . E . R) لتكشف عن أحسن الطرق ـ والوسائل لتطويع هذه النظرية للتطبيق .

ان تطبيق هذا النموذج في مجال التربية أو علم النفس يحتاج الى فريق من العلماء والباحثين ومسانـدة من هيئـات علمية سواء بالامـكانيات العلمية أو المالية . أن مقياس الـذكاء القومـي البريطاني الذي بدأ في عام ١٩٦٥ وأعلن أخيرا في نشرة مطبوعات N . F . E . R. انه سيصدر في خريف عام ١٩٧٨ قد قام بمساندة من لجنـة من الجمعية النفسية البريطـانية وجامعـة مانشستـر وكذلك المركز القومي للبحوث التربوية .

وقد تناول بناء المقياس فريقان من العلماء والباحثين ، بدأ الفريق الاول من عام ١٩٦٥ وحتى عام ١٩٧٣ ثم تسلم العمل فريق آخر وأكمل وطور عمل الفريق الاول . وقـد نفـذت ميزانية المشروع عام ١٩٧٠ ، ورصدت له ميزانية جديدة عام ١٩٧٣ .

لقد أرسى «جورج راش» عالم الرياضيات الدانماركي أسس هذاالنموذج ، ونما وتطور في الولايات المتحدة الامريكية ، وأستخدمته بلاد اخرى من بينها بريطانيا ، التـي أستعـان فريق

العمل بها بجهود ونصائح Ben Wright الاستاذ بجامعة شيكاغو وبامكانيات مركز الكمبيوتر بهذه الجامعة ، ذلك أن العلم تراث أنساني لا ينفرد به فرد أو أمة وهذا هو الانفتاح العلمي .

والآن أمامنا طريقان : أما أن ننتظر حتى تحل جميع مشاكل تطبيق النموذج ، ويصبح استخدامه أمرا سهلا من الوجهة العلمية ، ثم نأخذ النتيجة جاهزة وليس لنا سوى أن نفهمها ونطبقها بحذافيرها . وأما أن نساهم نحن أيضا في صنع التراث العلمي للانسان ، ويلزمنا عندئذ العمل كفريق مع مساندة علمية ومادية ، كما يلزمنا أيضا المتخصصون في علوم الكمبيوتر وتطبيقها في المجال السلوكي بصورة خاصة .

خاتمة

ترى الباحثة ان تعدد زوايا الرؤ ية في تفسير نتائج الاختبارت لا يحمل أية بلبلة على الاطلاق بل على النقيض فهي ترى في ذلك مزيدا من الاثراء لمعنى الدرجة الخام التي يحصل عليها الفرد على اختبار ما . وتدعو بوضوح الى وجهة النظر التي ترحب بفتح جميع النوافذ وتدعيم كافة الاتجاهات . واستخدام الاساليب والطرق المختلفة حتى يكون تأويل الدرجات أكثر معنى في خدمة كافة الاهداف وفي كل المجالات .

الحواشي

1 — Carver , R. P. Two Dimensions of Tests Psychometric and Edumetric. **American Psychologist.** . Vol. 29. N 7. July 1974.

٢ ـ المرجع السابق ، ص ٥١٢ .

3 — Papham , W. Criterian — **referenced Measurement, Educational Evaluation.** Ch. 7. Prentice—Hall. Inc. Englewood Cliffs. New Jersy. 1975.

4 — Murray, D., Rasch Item analysis and Scaling. In **the British Intelligence Scale**.Final report before Standarization. 1975 —1976. P. 470.

5 — Tyler, R.W. and Walf , R.W.A., Prespective on the issues. In Tyler, W., Walf, R.M. (eds). pp. 4 —6.

6 — Bloom, B B.S. Hastings, J.T., and Madans, F. Handbook of Formative and Evaluation of Student Learning. Mc Graw-Hill Book Co Inc. 1971. P. 45.

7 — Ebel, R.L. (1971), Criterions — referenced — Measurements Limitation, School review, 79, 282 — 288.

8 — Carver, R. D. Two Dimensions of Tests Psychometric and Edumetric. American Psychologist. Vol. 29. N 7. July 1974. P. 514.

٩ - المرجع السابق ص ٥١٤ - ٥١٦ .

10 — Ebel, R.L. (1971), Criterian — referenced — Measurements Limutation, School review, 79, 282 — 288

11 — Murray, D., Rasch Item analysis and Scaling. In the British Intelligence Scale. Final report before Standari zation. 1975 — 1976. P. 422.

١٢ - معصومة محمد كاظم . دور النماذج الرياضية في تطوير مفهوم الرياضيات التطبيقية في التعليم العام ١٩٧٨ .

13 — Willmot, S., and Fowles, D. The objective interpretation of test performance the Rasch Model applied. N.F.ERPublishing Co . Ltd . 1974 . P . 15 .

14 — Murray, D., Rasch Item analysis and Scating. In the British Intelligence Scale. Final report before Standarization. 1975 — 1976. P. 419.

١٥ - نفس المرجع السابق .

16 — Willmot, S. and Fawles, D. The objective interpretation of test performance , the Rasch Model applied. N.F.E.R. Publishing Co. Ltd. 1974 P. 15.

١٧ - المرجع السابق ص ١٦٠ .

١٨ - المرجع السابق ص ١٤ .

١٩ - المرجع السابق ص ١٩ -

٢٠ - المرجع السابق ص ٢٠ -

٢١ - المرجع السابق الملحق ، ص ٢١ .

٢٢ - المرجع السابق ص ٢٢ .

٢٣ - المرجع السابق .

24 — Murray, D., Rasch Item analysis and Scaling. In the British Intelligence Scale. Final report before standarization 1975 — 1976.

25 — Willmat, S,. and Fowles, D. The objective **interpretation of test performance**, the Rasch Model applied: N.F.E.R. Publishing co. Ltd. 1974, PP. 25 — 45.

26 — Murray , D. , Rasch Item analysis and Scaling. In The British Intelligence Scale. Final report before Standarization. 1975 — 1976. P. 429.

٢٧ ـ المرجع السابق ص ص ٤٢٩ ـ ٤٣٠ .

28 — Willmat , S., and Fowles, D. The Objective Interpretation of test performance, the Rasch Model Applied. N.F.E.R. Publishing Co. Ltd. 1974, PP. 34 — 37.

٢٩ ـ المرجع السابق ص ٣٩ .

